VIDEO ATTENTION ON SOCIAL MEDIA

A THESIS

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE

MASTER OF ARTS

BY

BREANNA L. DAUGHERTY

DR. ROBIN BLOM – ADVISOR

BALL STATE UNIVERSITY

MUNCIE, INDIANA

MAY 2019

Table of Contents

ACKNOWLEDGEMENTS

I would like to thank my academic mentor, advisor, and thesis committee chair Dr. Robin Blom for his continuous effort and assistance with my thesis. Additional acknowledgements go to Dr. Martin Smith-Rodden and Dr. Tim Huang for their tremendous support in my work. I would also like to extend my acknowledgements to members of the faculty and staff in the Department of Journalism, Dr. Renee Clear, Ryan Sparrow, and Jerry Pierce, for their help in making my thesis possible.

Video attention on social media

Introduction

Video has become a commodity in the digital world with videos being shared multiple times a day between friends and family. Over time, shared videos of funny advertisements, animals, food, and everything in between, have demonstrated society's increased interest in videos on social media platforms. One way to make the constant stream of videos engaging and attention keeping is to minimize the length of the video to under two minutes (Fishman, 2016). It has been speculated that the two minute mark has to do with humans' short attention spans, especially after the Internet became popular (Fishman, 2016). This study sought out information on what viewers attended to when there was a combination of a person, text (subtitles), and location (noise). It also examined if impression-relevant involvement paired with attitude, interest, and or recycling habits has an effect on an individual's attention.

With social media's growing prevalence in society's everyday life, it is important to understand what kind of video content consumers are more likely to watch, whether it is 10 seconds or two minutes. This information would be valuable to companies that produce sponsored advertisements and news organizations trying to inform the public. Being able to produce video content the consumers want to see could increase profits for companies, and better inform the public. Kumar and Gupta (2016) suggested research on making the content more relevant, targeted, and timely. Additionally, there has been minimal research focused on video content and its relationship with social media, which would be very beneficial as social media continues to develop. The present study could aid in the development of future research in this emerging field.

In particular, it is important to understand more about what content consumers are looking at when they are on social media. When organizations, marketers, and companies alike have a better understanding as to what, and why, do individuals watch videos on social media, there can be more tailored content. Social media are constantly evolving, whether it is because of new platforms or new algorithms; which has made it challenging to research; the effect of video attributes to increase audience attention. As social media began to consume more of our daily lives, it is important to have a better grasp of the different platforms and what individuals are attending to. An important aspect about this research is it considers where people could be viewing social media. Anyone can look around and notice that people are typically looking at their phone, and this study took that into account. Furthermore, there has been a lack of research that considers social media usage and an individual's location, which could be an important reason as to why individuals give their attention to certain topics or posts on social media.

The design of this study was a 2x2x2 experiment, which looked at the presence of person (whether or not there was person present in the video), subtitles (whether or not subtitles were included in the video), and noise (whether or not there was loud background noise when viewing social media). The element of a person was used because of face attention modeling and mere exposure effect by Ma, Hua, Lu, and Zhang (2005) which stated that having a face within a video can attract an individual's attention, but also because of the familiarity of the face; as well as fluency theory which specified that people prefer visuals that are processed easily (Palmer, Schloss, & Sammartino, 2013). Both Ma et al. (2005) and Chen et al. (2003) have discussed that text in a video can play a role in an individual's attention. For this reason, and because of the noticeable popularity of subtitles, it was added an element in this study. Finally, noise was studied because people are not always viewing social media in quiet location, but also because

Ma et al. (2005) discussed aural saliency modeling which stated that the loud or sudden sound can attract a human's attention. While these elements have been discussed together, they have yet to be put into an experiment to be studied in the context of social media.

The current study also examined the extent to which attention to a video on social media comes from having a level of impression-relevant involvement paired with attitude and if interest plays a role. There has been a lack of research specifically focusing on the interaction between these items, as well as using eye-tracking technology. For the purposes of this study, the topic of the video was environmental friendliness, because it is a current topic of interest among Millennials, as reported. This generation has sought out products that try to have a positive impact on the environment, as well as purchasing products that are environmentally friendly (K. T. Smith, 2010). The present study included a student sample where participants with between the ages of 18 and 26 years from a midsize university in the Midwest. The results from the study included statistical significance in regards to greater attention given to subtitles, as well as attention to subtitles on a video when there is no noise. Specifically, this indicated that including subtitles in a video could lead to an increase in attention. Having this information can inform marketers, organizations, and companies of the possibility of increased attention; however, this could also be the focus of future research.

Literature review

Social media are becoming more common among the general population, as it has become easily accessible with handheld technology. Additionally, they have become important to make images and other content more adaptive to the different pieces of technology as the Internet presents visual content (Chen et al., 2003). Kaplan and Haenlein (2010) defined social media as "a group of Internet-based applications that build on the ideological and technological

foundations of Web 2.0, and that allow the creation and exchange of User Generated Content" (p. 60). Web 2.0 referred to description of a new way to utilize the World Wide Web as a platform where content was no longer created by individuals, but rather altered by all users in a participatory and collaborative fashion (Kaplan & Haenlein, 2010).

A. Smith and Anderson (2018) of the Pew Research Center reported that 80% of people between the ages of 18–24 use Facebook on a daily basis. Therefore, the sample was focused on individuals 18–26 years old because that is the typical age range for college students. This information leads the way as the most popular social media platform. Additionally, the Pew Research Center reported that individuals between the ages of 18–29 use social media more other age groups (Center, 2018); which accounts for the age range in the sample size. The Pew Research Center has also reported 74% of Facebook users check the social network daily, which was higher than any other social media platform (Center, 2018).

Facebook went from 5.5 million users in 2005 to 500 million users in 2011, allowing its popularity to flourish in the digital world (Hughes, Rowe, Batey, & Lee, 2012). Users can create a profile to post information about them ranging from their education, to their religious and political views, to their relationship status, and more. Additionally, both the user and their "friends" can share links, photos, videos, as well as other posts by users (Hughes et al., 2012). On Facebook, pictures accounted for 93% of the most engaging posts than other types of posts—text, videos, and text-based updates (Cooper, 2013), which indicated that people are more likely to engage with visual based posts.. Hughes et al. (2012) pointed out that photos, videos, and posts could also be created within group and fan pages on Facebook, which have become popular for online socializing. Facebook also has a feature that allows users to send messages to one another, much like text messaging (Hughes et al., 2012). Hughes et al. (2012) said, "Although

Facebook is the largest SNS, there are others. All social networking sites facilitate online, social interaction, yet they do not all offer the exact same services or have the same focus" (p. 561). For the purposes of this study, the social media platform that was utilized in the stimulus was Facebook.

The user attention model was utilized in this study to understand how the attention from the user can play role with a video, as well as for the definitions of attention and elements. The user attention model is an estimation of how much attention viewers may give to video contents when they watch a video (Ma et al., 2005). A video has been defined as being a combination of "image sequence, audio tracks, and textural information" (Ma et al., 2005). Each element of the video delivers information to the viewers, according to Ma et al. (2005). The image sequence was composed of motion (object motion and camera motion), appearance (color, shape, texture), and objects (Ma et al., 2005). The audio tracks include elements of speech, music, salience, and various special sounds (Ma et al., 2005). The textual information was represented in linguistic form, such as closed captions, automatic speech recognition (ASR), and superimposed texts (Ma et al., 2005). Additionally, videos use guided attentions embedded in the video; such as camera motion and how loud sounds are in a video.

Within the user attention model, there is the motion attention model, which states that viewer attentions are caused by motion and are estimated based on motion vector field (MVF) (Ma et al., 2005). The motion vector field helps measure motion attention by using motion vectors. These vectors can be viewed as the "perceptual responses of the optic nerves" (Ma et al., 2005, p. 9). Ma et al. (2005) stated that motion with high intensity attracts more human attentions; though, camera motions also contribute to motion intensity. On the other hand, there is the static attention model; which can attract human attention; however, it cannot be detected or

estimated by the motion attention model (Ma et al., 2005). Next, there is aural attention modeling, or more specifically, the aural saliency model. This model looks at the essential assets of sound, which fundamentally include loudness (Ma et al., 2005). This is because "humans are often attracted by loud or sudden sounds if they have no subjective intentions" (Ma et al., 2005, p. 14). Ma et al. (2005) modeled aural saliency when there is either "an absolute loud sound, which can be measured by average energy of sound" or "the sudden increases or deceases of the loudness" (p. 14). Furthermore, Vachon, Hughes, and Jones (2012) discussed how one's auditory sequence allows inferences (or expectancies) to be consequential; however, unexpected auditory stimulus may be distracting because of its novelty.

Additionally, text regions also tend to attract attention by viewers, specifically if the text is larger than 10 points, which is referred to as text attention modeling (Chen et al., 2003). These models are frameworks and how they should work within the experiment and research. One aspect to attract human attention is by having the appearance of faces within a video, also referred to as face attention modeling (Ma et al., 2005). Because of these viewpoints, Ma et al. (2005) has purposed that both people and text play some type of role in video attention. It is worth mentioning that there are attention models focusing on audio, which focuses on speech and music (Ma et al., 2005).

This purposed idea by Ma et al. (2005) could be related to fluency theory. Fluency theory postulated that people prefer visuals that can be processed more easily or fluently (Palmer et al., 2013). This theory provided a reasonable explanation as to why humans may be more interested in humans—often in media content—because there is a level of familiarity, so it is easier to process. This same ideal helps form the mere exposure effect. Mere exposure effect is caused by seeing an image more frequently which in turn, allows it to be processed more easily and fluently

(Palmer et al., 2013). As theorized by Zajonc (1968), the higher frequency of seeing an image or object increases preference to that particular image or object (see also Palmer et al., 2013).

Stroud (2017) described the ability to allocate attention as hard-wired into humans; however, he noted that it's not possible to pay attention to everything in our lives. "Rather, we allocate our attention to particular stimuli in the environment and ignore everything else," (Stroud, 2017, p. 480). For example, marketers are trying to reach their consumers on social media. Kumar and Gupta (2016) discussed how marketers solicited and customized advertisements specifically for social media, because that is how they want to reach their audience and get consumers to engage with a brand. This is important to note because it shows that people are on social media more than they were in the past. Moreover, with social media becoming more popular, and developing relevant content and keeping the consumers' attention span has become a challenge (Kumar & Gupta, 2016).

When thinking about attention and content types, there are two aspects: negative and positive. In everyday life, the negative, or bad, events have a great power over the positive, or good, events (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). This is because negative information is processed "more thoroughly" than the positive information (Baumeister et al., 2001). Bad has been defined as "undesirable, harmful, or unpleasant" (Baumeister et al., 2001, p. 326). Baumeister et al. (2001) defined "good" as "desirable, beneficial, or pleasant outcomes including states or consequences" (p. 325). While Baumeister et al. (2001) said the power of positive can outweigh the negative, and has come close to the reversal with the anticipation of future events. For example, when people make a future decision, "they seem more motivated to avoid bad outcomes than to pursue good ones" (Baumeister et al., 2001, p. 356). However, "good can only match or overcome bad by strength of numbers" (Baumeister et al., 2001, p. 356).

Alternatively, Baumeister et al. (2001) found that negative is stronger than positive in health, social support, and learning topics. An additional reason as to why humans may process negative information more thoroughly than positive is because of our evolutionary need that signals a change should happen (Baumeister et al., 2001). On the other hand, according to Baumeister et al. (2001), "good may prevail over bad," more specifically he references psychological events. Despite research and determination by Baumeister et al. (2001) that bad is stronger than good, for the purposes of this study, the video stimulus that will be used will be set in a positive tone. This is with the idea of by Baumeister et al. (2001) that when people make a future decision they wish to avoid bad outcomes.

Attention has often been referred to as a mental effort and selective processing (Johnston & Dark, 1986; Johnston & Heinz, 1978; Posner & Boies, 1971). Warshaw (1978) said, "Broadbent's theory of selective attention implies that the perception may not obtain unless audio and video are properly sequenced" (p. 366). This was because Broadbent (1958) suggested that a limited capacity filter between the recognition and attentive stages of analysis restricts information processing by humans. Additionally, Warshaw (1978) found evidence that proposed it takes time to switch between attention channels; including Broadbent (1971) that claimed it took about 1.5 seconds. Nonetheless, Broadbent's examination finds that only one channel between audio and visual channels are given attention (Warshaw, 1978). "Selective attention refers to the differential processing of simultaneous sources of information" (Johnston & Dark, 1986, p. 44). In relation to selective attention theory, there is "biased competition theory," proposed by Desimone and Duncan (1995), that illustrates attention as a signal competition within the brain (Webb & Graziano, 2015). This was because the signals compete which will be processed (Webb & Graziano, 2015). According to Webb and Graziano (2015), the brain has to

"control attention in a task-relevant fashion" (p. 4). As noted by Johnston and Dark (1986), more research has used the top-down processing for selective attention, which focuses on what the individual is interested within the stimuli. Alternatively, there was bottom-up processing with selective attention, which acts as a data-driven processing system where "stimulus activates codes at various levels of analysis ranging from physical or sensory analysis to complex semantic and schematic analysis" (Johnston & Dark, 1986, p. 44).

Giving one's attention to a video also encompasses being involved with what is being shown. Cho and Boster (2005) wrote that development on involvement studies has provided "pivotal constructs" in understanding messages and media influence audience behavior in the communication field. In their study, Cho and Boster (2005) focus on value-, outcome-, and impression-relevant involvement. For the purposes of this study, impression-relevant involvement, which concentrated on the "concern with public perception of the self" (Cho & Boster, 2005, p. 239), was utilized. This was because value-relevance involvement refers to one's ego, which showed "the characteristic feelings of continuity and permanence the individual has about himself" (Sherif & Cantril, 1947, p. 94). In addition, the outcome-relevant involvement was thought to be too vague by focusing on the subject of an individual's goals or outcomes (Cho & Boster, 2005).

Differing from the aforementioned value- and outcome-relevant involvements is the impression-relevant involvement. Zimbardo (1960) described this as an "individual's concern with the consequences of his response or with the instrumental meaning of his opinion" (p. 87), as stated by Cho and Boster (2005). From Zimbardo's experimental inductions, Cho and Boster (2005) wrote that the only consequence is an individual's impression on others. Furthermore, impression- and outcome-relevant involvements differ because of their task orientations (Cho &

Boster, 2005; Leippe & Elkin, 1987). Impression-relevant involvement prompts a public-task orientation because individuals have a concern with others' perceptions (Cho & Boster, 2005). Consequently, outcome-relevant involvement motivates private-task orientation because individuals are inspired to meet personal goals and standards (Cho & Boster, 2005). Cho and Boster (2005) wrote that impression-relevant involvement is positively associated with the dimension of the self-monitoring scale; which looks at how self-monitoring relates to a concern that orientated toward people by "pleasing others, conforming to the social situation, and masking one's true feelings" (Briggs, Cheek, & Buss, 1980, p. 681; see also Cho & Boster, 2005, p. 240). Each of these involvement types plays into attention and how it could be divided among the current study.

Because there has been a lack of research regarding impression-relevant involvement specifically focuses on environmental friendliness, the issue of environmental attitude was added. Environmental attitudes, as Schlegelmilch, Bohlen, and Diamantopoulos (1996, p. 38) are defined as "capturing individuals' levels of concern/interest about specific or general aspects of environmental, ecological, or energy-saving phenomena" (see also Buttel, 1979). Because impression-relevant involvement looks at an "individual's concern with the consequences of his response" (Cho & Boster, 2005, p. 85) and the attitudes look at an "individuals' levels of concern/interest about specific or general aspects of environmental" (Schlegelmilch et al., 1996, p. 38). These similarities in concepts allow the opportunity to look at involvement with attitude when discussing environmental friendlessness.

To add an additional reason for one's attention, and possible involvement, in a video could be due to the psychology of interest. The psychology of interest has been theorized into two fields: emotion and personality. The emotional part of interest involves an emotional

experience, curiosity, and momentary motivation (Silvia, 2006). Interest and emotion, or "interest," have been recognized when momentary feelings of interest have been studied with context, such as when individuals find paintings, music, and texts interesting. Other researchers have found additional ways to recognize interest and emotion with facial and vocal expressions and motivational effects of interest, which have been examined as causes and consequences of feeling interested. While the other field interest focused on personality, individual differences and idiosyncratic hobbies, goals, and avocations of people (Silvia, 2006). Additionally, aspects of interest and personality have been studied with context that found traits related to curiosity, openness to experience, sensation seeking, boredom and breadth of interest (Silvia, 2006). Silvia (2006) refers to interest with emotion as "interest" and interest with personality as "interests," for simplicity, we will do the same in this study. In relation to one's involvement, both interest and interests could play a role in the reason why an individual finds visual content interesting. However, for the purposes of this study, we will only be looking to the interest of individuals. This is because Silvia (2006) wrote "interest motivates exploration and learning, and guarantees the person's engagement in the environment" (see also Izard, 2000). It is with the ideal of "momentary motivation" that interest was used within this study.

The involvements, attitudes, and interest that could be pertinent to an individual in this study regard the environmentally friendly movement. Lai, Cheng, and Tang (2010) found that "green retailing" is common occurrence and is viewed as an inclusion measure to help protect the environment within retail procedures. Furthermore, K. T. Smith (2010) discussed a California Green Solutions' 2007 report that stated 47% of Millennials would pay more for environmentally friendly products, brands, and services. In addition, 77% of them said they would do it because they "care about the environment." K. T. Smith (2010) also cited Gunelius' 2008 report that

Millennials seek out brands that trying to make a positive impact on the environment, and found the most effective words to use to convey to Millennials a product is environmentally friendly are: “eco-friendly, recycled, and green” (p.10). The study also found that Millennials perceive the recycling symbol as synonymous with being environmentally friendly, in fact 90% of the Millennials studied felt that way. When asked what would motivate Millennials to promote a product to their friends, they suggested green products that are economical and products made of recycled materials. This information helped determine what content will be used in the video stimulus portion of the study. When videos were being considered for stimulus in the study, they had to discuss the topic of being “eco-friendly, recycled, and or green,” which are keywords, according to K. T. Smith (2010).

Hypotheses development

There has been a noticeable lack of research that looks at face attention modeling and fluency theory, as well as the use of subtitles; even more specifically they possibly have not been studied together. According to Ma et al. (2005), having faces appear in a video can aide in human attention, known as face attention modeling. This information paired with fluency theory—which states that people prefer images they process easier, which also causes the mere exposure effect (Palmer et al., 2013)—furthers the idea that humans prefer to look at humans when viewing visuals. As Ma et al. (2005) stated “the appearance of dominant faces in video usually attracts viewer attentions” (p.12). Using face attention modeling fluency theory, and mere exposure effect, the following hypothesis was developed:

> H1a: Attention will be higher when a video contains a person talking than a video that does not contain a person.

There has been a lack of conversation in previous literature regarding textual elements within a video, specifically on social media, and how users attend to these elements. It is important to consider attention on textual elements in videos, as subtitles have become increasingly popular among social media users. Perego, Del Missier, Porta, and Mosconi (2010) discussed how research on subtitles could have value because of the lack of focused research on the subject. Other ways subtitles have become more popular are to support those that are deaf or hard hearing, to promote literacy, a low cost way translation for films, and more (Perego et al., 2010). Additionally, Chen et al. (2003) said viewers appear to give attention to textual elements. Because of the multi-purposes subtitles can be used for, it is important to look at how attention on subtitles in a video are being attended to or at all.

H1b: Attention will be higher when a video contains subtitles than a video that does not contain subtitles.

While there has been previous research that examines how attention can deter a person's attention (Vachon et al., 2012), there has been a gap of research that specifically looks at social media, video, and noise together. However, it can be hypothesized using an ideal from Ma et al. (2005), which discusses how aural information can be an important source of human attention. Taking this into consideration, aural attention could be used at the time of looking at social media in a noisy area.

H1c: Attention will be lower when participants watch a video in a noisy area than a video that is not in an area without noise.

This study is not directly looking at attention and sound relation. Sound can play a role in attention from aural saliency modeling, according to Ma et al. (2005). However, with recent updates to Facebook and other social media platforms, videos do not automatically have the

sound on when they start playing as you scroll across a page. Depending on the video, there will subtitles until an individual clicks on the volume button or video to allow sound on the social media platform. Because there has been a lack of research in the area of audio and video attention, it is posed as a research question to see what participants do when they are exposed to a video on social media.

RQ1: Will individuals want to listen to the audio of video when there are not subtitles?

An additional concept to take into consideration is the location of an individual—if it is a noisy or quiet area—and how it will affect their attention to a video—especially when text is involved. It is important to consider topics, such as text, when viewing social media, because there has been a lack of research to determine when or if an individual will turn on their sound while being in public. Considering the possibility of individuals watching videos in public crowded areas, the following hypotheses were developed:

H2a: For an individual in a noisy area, attention will be higher when a video contains subtitles than a video that does not contain subtitles.

H2b: For an individual in a noisy area, attention will be higher when a video contains a person talking, unless there are no subtitles.

Both Ma et al. (2005) and Chen et al. (2003) have stated that text can play a role in an individual's attention when looking at a video. Furthermore, there was the lack of noise, which eliminates the possibilities of aural attention modeling, which states that individual's could be distracted by outside noise (Ma et al., 2005). Considering these models, the following hypothesis was developed:

H2c: For an individual in an area without noise, attention will be higher when a video contains subtitles, than a video where there are no subtitles.

Now, using the ideals of fluency theory, face attention modeling, and mere exposure effect which state people allocate more attention to humans, because it's easier to process humans as individual's are familiar with faces (Palmer et al., 2013), as well as having faces, and the familiarity of the faces, can attract an individual's attention (Ma et al., 2005). Again, there is no direct aural distraction, which are the opposite effects of aural attention modeling. Using fluency theory and face attention modeling with the opposite effects of aural attention modeling, the following hypothesis was developed:

H2d: For an individual in an area without noise, attention will be higher when a video contains a person, than a video without a person.

Taking into consideration, once more, aural attention modeling by Ma et al. (2005) looked into how an individual's attention may focus on where noise is coming from. For example, in the aural saliency modeling, Ma et al. (2005) stated that "human are often attracted by loud or sudden sounds if they have no subjective intentions" (p 14). Because individuals might be distracted by the noise in an area, it can be hypothesized that they would move more quickly through a social media newsfeed.

H3: For an individual within a noisy area, they will arrive to the video more quickly than an individual that is not in a noisy area.

Taking the aural saliency model (Ma et al., 2005) into consideration when it comes to attention, it can be hypothesized that an individual's may be divided when in a noisy area. Much like the aforementioned hypothesis, the aural saliency model stated that outside noises can deter one's attention, especially when there are loud or sudden sounds (Ma et al., 2005). With this information available, the following hypothesis was developed:

H4: Attention to subtitles will be higher in an area without noise than in a noisy area.

As Ma et al. (2005) and Chen et al. (2003) stated, text can play a type of role in an individual's attention when pertaining to a video. Furthermore, Ma et al. (2005) and Palmer et al. (2013) discussed how face attention modeling, fluency theory, and the mere exposure effect can cause individuals to pay more attention to a particular video because there is a person in it, which was also familiar. The following hypotheses were developed with considering only the subtitle condition:

H5a: When there are subtitles in a video, attention will be higher when a video containing a person is played in an area without noise than when a video is played in an area with noise.

H5b: When there are subtitles in a video, attention will be higher when a video containing no person is played in an area without noise than when a video is played in an area with noise.

Method

Experimental design

This study presented an experiment that was a 2 (person, no person) x2 (subtitles, no subtitles) x2 (noise, no noise) design. This required 8 groups to be created within the experiment, alternating person, subtitles, and noise, in which two or three factors were used for each participant. Taking fluency theory and face attention modeling into consideration, the person, no person condition was added because of an individual's innate ability to focus on a human, according to Palmer et al. (2013). Because of the aforementioned ideals by Ma et al. (2005) and Chen et al. (2003), text in a video may cause an individual to attend to it. Moreover, subtitles are becoming an increasingly popular addition to videos. Finally, when considering how individuals behave on social media, it is fairly obvious to assume that society, as a whole, looks at different

platforms while being in public. Aural saliency attention modeling by Ma et al. (2005), paired with this knowledge, lead to the noise, no noise condition. Additionally, the dependent variables in this study were total fixation and total fixation subtitles because of the different combinations of the design could influence total. The environmental attitude scale, environmental interest scale, and if a participants recycles to be the control variables, in which it was examined if these factors had an impact on the results.

Stimulus and Facebook

For the purpose of this study, a video was used that focused on recycling. The video, by EcoProducts, was chosen because of the previously mentioned research regarding Millennials interest in environmentally friendly practices. One of the ways Millenials recognize a product as environmentally friendly is by using a term like "recycled" (K. T. Smith, 2010). EcoProducts' video discussed the recycling loop and how they're working to close it by working with local restaurants and recycling centers a particular area. Additionally, this video contained a positive tone, rather than a negative one, using an ideal from Baumeister et al. (2001) that individuals will want to prevent bad outcomes when making future decisions. The videos were posted on the researcher's YouTube channel, and were uploaded as unlisted; this allowed the videos not to appear on the researcher's YouTube channel. Additionally, for the duration of the study, the researcher changed their YouTube profile picture to that of the EcoProducts logo, as to not reveal the authenticity of the video.

After reviewing a little over 20 companies' YouTube and Facebook accounts for videos—some including The Coca-Cola Company, Patagonia, Unilever, and Nike—EcoProducts had a video that contained the necessary pieces for the research. Some of those pieces included having a person speaking to the viewer, adequate amount of B-roll, and enough footage to work

with for editing purposes, and other videos on their account to use for additional B-roll. B-roll was defined as "extra footage captured to enrich the story you're telling and to have greater flexibility when editing" (Jellinek, 2011). Other videos may have provided a person, but not enough for B-roll for editing purposes. After the video was selected, it was downloaded from YouTube and put into Adobe Premiere to be edited. The video was edited down below 2 minutes, in all four versions of the video— no person, no subtitles; no person, subtitles; person, no subtitles; and person, subtitles (see examples of the beginning of the videos in Appendix A). Additionally, the videos with no person were edited to remove the person talking to the viewer. Each video carried the same message and idea, but was adjusted to fit the appropriate condition. Additionally, subtitles were entered in by using the open captions function on Adobe Premiere using Helvetica 19 point font, because Chen et al. (2003) noted that text had to be at least 10 points to attract attention.

In the video, a member of EcoProducts discussed how their products are used at an eatery in Denver, Colorado and how they can be recycled. Snooze, the eatery in Denver, has their sustainably maven discussing their mission of being able to recycle with EcoProducts and having a goal of being a zero waste kitchen. Then, the sustainability maven discussed their recycling process with a local recycling center, which led to another individual speaking about the recycling center.

Previously mentioned, participants were exposed to one of the four videos, but they were also exposed to an additional element of noise. This noise was used to determine if outside noise plays a role in attention when viewing videos on social media. There has been a lack of research in this area, which is why it was tested in this study. The noise participants were exposed to was the sound from a video made in Times Square in New York City, New York. The video

represented what it was like to be in a highly congested area with people walking by and talking, cars driving and honking their horns, as well as general city noises (PollyPoly, 2017). The noise provided in the video follows aural saliency modeling, according to Ma et al. (2005), as the sound may suddenly increase or decrease in loudness.

Facebook

Because this study examined how individuals use social media, a Facebook emulation newsfeed was designed as part of the stimulus. For several days in December 2018, posts were observed from Facebook newsfeeds to determine what kind of posts would be in the final design. This included topics about having a child, an NPR news story, sports, a local business post, a travel related status, and changing a profile picture; as well as the video and post by EcoProducts. Additionally, names were created using a fake name generator; however, there is no way to ensure that these are not the names of real people. There were a total of 9 posts in the design (see full design in Appendix B).

After designing the newsfeed in Adobe InDesign, the file was exported for HTML using the Ajar in5 plugin for Adobe. Then, Brackets, a text editor, was used to finalize the design, as well as code the video into the HTML code. This code was used to ensure the video would begin playing when the participant scrolled to it on the newsfeed, as it typically behaves on Facebook. This allowed the design to be a webpage so users were able to scroll more naturally, as they would on any other desktop. Additionally, users had the display in full screen mode, as to not reveal the authenticity of the Facebook newsfeed by being able to view the search bar. The sound on the computer was also muted because sound does not automatically play when viewing content on a user's phone until the video is opened or the user turns on sound. When marketing to consumers, Mintel found that 72% of adults, 18 years and older, are more likely to trust claims

by a product if they say they are recyclable in an easy to understand way (O'Donnell, 2013) Therefore, the text within the post paired with the video by EcoProducts also focused on a message of recycling.

Survey

The person variable primarily looked at if there was a person talking to the viewer when it started playing; thus using fluency theory and face attention modeling (Ma et al., 2005; Palmer et al., 2013). Subtitles used the discussion by Ma et al. (2005) and Chen et al. (2003) on how text could cause an individual's attention to be focused on text, and was determined by the use of subtitles in a video. Finally, the noise variable considers the possibility of an individual being in a noisy area, rather than a quiet one, which uses the ideals of aural saliency modeling (Ma et al., 2005).

Overall, the survey, which was delivered at the end of the experiment portion, was aimed for measuring the control variables—environmental attitude scale, environmental interest scale, and recycling—that could play a considerable role in further explaining fixation duration beyond the experimental manipulations (see the full survey in Appendix C). The impression-relevant involvement and attitude statements regarding one's environmentally friendliness was on a 5-point Likert scale, ranging from strongly disagree to strongly agree. As previously mentioned, characteristics of impression-relevant involvement could play an additional role in an individual's attention to a video regarding environmentally friendliness. There was a lack of research focusing on impression-relevant involvement when focusing on environmental issues, therefore the researcher developed some of the statements on this variable. Additional statements were used from Bohlen, Schlegelmilch, and Diamantopoulos (1993) regarding environmental attitude statements. These statements were added because supplemental items were needed to

account for the lack of previously researched environmental scales using impression-relevant involvement; therefore, leading to the use of an environmental attitude scale. This scale was used as a partnership with impression-relevant involvement. Both look at an individual's "concern," which applies to both in the way they can be perceived. While the data from the survey did not compound with all the statements, one scale—the environmental attitude scale—was utilized in this study ($M = 4.1$, $SD = 0.6$). A confirmatory factor analysis of the items was conducted using varimax rotation with one factor explaining 59% of the variance. All items in this analysis had primary loadings over .7. Internal consistency for the scales was examined using Cronbach's alpha (.86).

Participants were then asked to respond to statements regarding their interest regarding an individual's interest level in recycling at the end of the experiment by using a 5-point Likert scale, which ranged from strongly disagree to strongly agree. These statements were used because an individual could already be taking part in the environmentally friendly movement; therefore, altering their attention. There was lack of research focusing only on environmentally friendly interest levels. Because of this, statements were developed to measure interest. Not all of the data from the survey compounded, therefore one scale—the environmental interest scale—was utilized in this study ($M = 4.1$, $SD = 0.5$). A confirmatory factor analysis of the items was conducted using varimax rotation with one factor explaining 55% of the variance. All items in this analysis had primary loadings over .7. Internal consistency for the scales was examined using Cronbach's alpha (.73).

Furthermore, a single variable that was measured was whether or not if individual's recycled. This exploratory variable was developed by the researcher to help determine whether or not recycling habits play a role in an individuals' attention to recycling stimuli. This was a

dichotomous question that simply asked “Do you recycle?” where participants answered “yes” or “no.” Of all participants ($N = 95$), 85% indicated that they recycle ($SD = 0.36$).

Sample

Participants in this study were a convenient sample ages ranging from 18 to 26 years old ($M = 19.7$, $SD = 1.4$) from a midsize university in the Midwest Almost three quarters of the participants identified as female ($n = 69$), and just over a quarter identified as male ($n = 25$); and one person preferred not to answer. It is important to note that in this study, there is “a natural confounding” variable regarding personal relevance and an issue could be the amount of thinking a person has done about the issue before taking part in the experiment (Petty & Cacioppo, 1986). Participants for the experiment were divided into groups for each condition to ensure group equivalence. Each condition had a combination of the independent variables: (1) no person, no subtitles, no noise; (2) no person, no subtitles, noise; (3) no person, subtitles, no noise; (4) no person, subtitles, noise; (5) person, no subtitles, no noise; (6) person, no subtitles, noise; (7) person, subtitles, no noise; or (8) person, subtitles, noise.

Data collection

Eye tracking technology was used to analyze eye movements in terms of fixations—pauses over regions of interest—and saccades—rapid movements between fixations (Salvucci & Goldberg, 2000). Some common analysis metrics, as listed by Salvucci and Goldberg (2000), are fixation or gaze durations, saccadic amplitudes, saccadic velocities, and transition parameters between fixations and/or regions of interest. To analyze fixations and saccades, there has to be a form of fixation identification that is the translation of raw eye movement data points to fixation locations on the visual display (Salvucci & Goldberg, 2000). The identification removes the raw saccade data points and collapses raw fixation points into a single representative tuple (Salvucci

& Goldberg, 2000). While Salvucci and Goldberg (2000) noted that it was agreed upon that visual and cognitive processing take place during a fixation, it is less clear when the fixation starts and ends. For the purposes of this study, total fixation duration and total fixation subtitle duration were utilized. An additional reason why eye-tracking technology was used is because it is a validated method of measuring attention via the corneal reflection in eye tracking (Vraga, Bode, & Troller-Renfree, 2016). For the purposes of this study, the Tobii eye tracker X2-30 Compact with Tobii Studio 3.4.8.1348 was employed.

Within the Tobii Studio 3.4.8.1348 software, heat maps were generated as a way to understand where participants are directly looking, while reviewing the stimulus (Bojko, 2009). Heat maps are valuable because they can summarize data of large quantities, as well as understand various types of data, "such as usage, accuracy, or visual attention" (Bojko, 2009, p. 1). Essentially, each entry on a heat map is from raw data that corresponds to a raw gaze point from the eye tracker. Heat maps were utilized in the present study to allow a better understanding of what participants are attending to.

Procedure

The researcher recruited students in eight journalism and telecommunications classes. The students were offered 2% extra credit for their participation, in which they signed up for an appointment to take part in the study via a Doodle Poll. Before each session, the researcher prepared the appropriate condition. Participants signed a consent form before beginning the study, which also allowed them to receive the extra credit in the class that they received the recruitment message.

Participants began their session with by having their eyes calibrated so the Tobii X2-30 eye tracker would properly record their eye movements. This required the researcher to manually

calibration each individual participant; afterward, participants began looking at the newsfeed. Because the newsfeed was created in HTML code on a website, the method of recording the participants' session was to use the screen recording feature on Tobii Studio 3.4.8.1348. This required that the researcher be present to exit out of the screen recording when the participant verbally acknowledged to the researcher they were finished viewing the newsfeed. Additionally, the participants in the researcher manually turned the sound on and off for participants in the noise conditions. When the survey portion began, the researcher moved away from the participant, so the participant did not feel pressured to answer a particular way. When they were finished with their 10–15 minute session, a debriefing screen appeared and the participants had completed their session (see the full procedure in Appendix D).

Analysis

Before transferring the data in to SPSS, Areas of Interest (AOI) had to be created in the Tobii Studio 3.4.8.1348 software. Tobii (n.d.) defines AOIs as a "tool that allows the eye tracking researcher or analyst to calculate quantitative eye movement measures" (para. 2), as well as fixation counts and durations. To put it simply, a tool is used to draw a boundary box around the element that is of interest—in this study, the video and subtitles. When making the AOIs, which are measured in pixels (Tobii, n.d.), the AOI around the video was approximately 520x227 pixels and the AOI around the subtitles was approximately 250x50 pixels. Furthermore, dynamic AOIs were used within in the Tobii Studio 3.4.8.1348 software because of the scrolling manner that took place during the experiment. These AOIs are defined by using keyframes in the timeline while creating the AOIs.

T tests were used to measure the data because *t* tests compare two means. This is because the means were assumed to be independent from each other, and because the populations were

almost even; only one group of the eight was missing a participant. Analysis of Covariance (ANCOVA), which is a blend of an analysis of variance (ANOVA) and a regression analysis, assessed the influence of the independent variables—person, subtitles, and noise—on the dependent variables (total fixation and total fixation subtitles) while controlling for the covariate factors. Two ANCOVAs, one that measured three independent variables and control variables, and the second that measured person and noise variables and control variables, were conducted because for the total video fixation the data for all participants was available. However, only half the sample was exposed to subtitles.

Results

Hypothesis 1a stated that attention would be higher when a video contains a person, compared to a video that does not contain a person. The average amount of fixation duration for participants in the human condition was 2.2 seconds (*SD* = 1.8), which was longer than the average time in the no person condition (M = 2.0, SD = 2.0). The independent sample *t* test indicated that this was not a statistically significant difference among the groups, $t(88) = -0.33$, $p = .74$. This indicated that there is no evidence in support of hypothesis 1a, although the mean difference was in the expected direction.

Hypothesis 1b stated that attention would be higher in a video that contains subtitles, than a video that does not contain subtitles. The average amount of fixation duration for participants in the subtitles condition was 1.7 seconds (*SD* = 1.4), which was significantly shorter than the average time in the no person condition (*M* = 2.6, *SD* = 2.2). The independent sample *t* test indicated that this was a statistically significant difference among the groups, $t(88) = 2.3$, $p = .02$. This indicated that there was evidence in support of hypothesis 1b (see Figure 1, Appendix E).

In this present study, hypothesis 1c stated that attention would be lower when participants are in a noisy area, compared to an area without noise. The average amount of fixation duration for participants in the noise condition was 1.9 seconds (*SD* = 1.6); that was shorter than the average time in the no person condition (*M* = 2.4, *SD* = 2.1). The independent sample *t* test indicated that this was not a statistically significant difference among the groups, $t(88) = 1.2$, $p = .22$. This signified that there was no statistical evidence in support of hypothesis 1c. However, the effect was in the expected direction.

These three mean comparisons did not control for the other factors—person, subtitles, and or noise. An ANCOVA measured the three factors, as well as three covariates—recycling, environmental attitude scale, and environmental interest scale—also indicated that there was a statistically significant role of subtitles in predicting total fixation, $F(1, 79) = 5.46$, $p = .02$ with a moderate effect size ($\eta_p^2 = .07$). The other independent variables and covariates were not statistically significant with low or even trivial effect sizes. There were no interaction effects among the three independent variables. The ANCOVA result was similar to the tests mentioned above.

Table 1

Total video fixation duration

Source	*Type III Sum of Squares*	*df*	*Mean Square*	*F*	*Sig.*	*Partial Eta Squared*
Corrected Model	45.563[a]	10	4.556	1.236	.232	.144
Intercept	26.889	1	26.889	7.823	.006	.090
Recycle	4.181	1	4.181	1.216	.273	.015
Environmental Attitude Scale	4.572	1	4.572	1.330	.252	.017

Environmental Interest Scale	.965	1	.965	.281	.598	.004
Person	1.378	1	1.378	.401	.528	.005
Noise	2.964	1	2.964	.862	.356	.011
Subtitles	18.779	1	18.779	5.463	.022	.065
Person * Noise	2.531	1	2.531	.736	.393	.009
Person * Subtitles	4.561	1	4.561	1.327	.253	.017
Noise * Subtitles	6.351	1	6.351	1.848	.178	.023
Person * Noise * Subtitles	1.279	1	1.279	.372	.544	.005
Error	271.548	79	3.437			
Total	719.320	90				
Corrected Total	317.110	89				

a. R = .144 (Adjusted R^2 = .035).

The research question asked in this study was if individuals would want to listen to the audio from the video if there are no subtitles. During the study, the computer was set to mute, because, as previously discussed, Facebook and other social media platforms do not automatically play sound when a video starts playing. It was observed that a few participants hovered their mouse over or near the sound icon on the YouTube video display, and when they realized the sound was on for the video, but the computer was muted, they didn't attempt to turn on the sound. However, one participant unmuted the computer by using the keyboard to watch a portion of the video before deciding to move on. This participant did not have the noise condition, but rather the person, no subtitles condition.

In this study, hypothesis 2a stated that an individual's attention would be higher in a noisy area when a video contains subtitles, compared to a video that does not contain subtitles. The average fixation duration for videos with subtitles (M = 1.7, SD = 1.4) was shorter than for

videos without subtitles (M = 2.0, SD= 1.8). The independent sample t test indicated that this was not a statistically significant difference, $t(44) = 0.72$, $p = .48$. This provided no evidence in support of hypothesis 2a.

Hypothesis 2b stated that an individual would have higher attention in a noisy area when a video contains a person, unless there are no subtitles. The average fixation duration for videos with a person (M = 2.1, SD = 1.6) was longer than for videos without a person (M = 1.7, SD= 1.6). The independent sample t test indicated that this was not a statistically significant difference, $t(44) = -0.76$, $p = .45$. This provided no evidence in support of hypothesis 2b in this study.

Hypothesis 2c stated that attention would be higher for an individual in an area without noise when a video contains subtitles compared to a video where there are no subtitles. The average fixation duration for videos with subtitles (M = 1.6, SD = 1.4) was shorter than for videos without subtitles (M = 3.1, SD= 2.5). The independent sample t test indicated that this was a statistically significant difference, $t(42) = 2.5$, $p = .02$. However, this did not provide evidence in support of hypothesis 2c, because the difference was in the opposite direction.

Hypothesis 2d stated that attention would be higher in an area without noise when the video contains a person than when the video is played in a noise area. The average fixation duration for videos with a person (M = 2.3, SD = 2.0) was slightly shorter than for videos without person (M = 2.4, SD = 2.4). The independent sample t test indicated that this was not a statistically significant difference, $t(42) = 0.17$, $p = .87$. This provided no evidence in support of hypothesis 2d.

Hypothesis 3 stated that an individual would have higher attention to a video with subtitles in an area without noise, compared to a video where there are no subtitles. The average

amount of time to first fixation to the AOI for participants in the noise condition was 33.0 seconds (*SD* = 17.5); that was shorter than the average time in the no person condition (*M* = 37.2, *SD* = 16.9). The independent sample *t* test indicated that this is not a statistically significant difference among the groups, $t(90) = 1.2$, $p = .24$. This indicated that there is no statistical evidence in support of hypothesis 3. However, the effect was in the expected direction.

Hypothesis 4 stated attention to subtitles would be higher in an area without noise than an area with noise. The average amount of fixation duration for participants in the no noise condition was 0.4 seconds (*SD* = .51); that was shorter than the average time in the noise condition (*M* = 0.56, *SD* = 0.74). The independent sample *t* test indicated that this was not a statistically significant difference among the groups, $t(42) = 0.94$, $p = .35$. This signified that there is no statistical evidence in support of hypothesis 4. However, the effect was in the expected direction.

In hypothesis 5a, it was stated that attention would be higher in a video containing subtitles and a person when it is an area without noise than area with noise. The average amount of subtitles fixation duration for participants in the no noise condition was 0.8 seconds (*SD* = 0.9); that was longer than the average time in the noise condition (*M* = 0.3, *SD* = 0.6). The independent sample *t* test indicated that this is not a statistically significant difference among the groups, $t(19) = 1.58$, $p = .13$. This indicates that there was no statistical evidence in support of hypothesis 5a, even though the time in the noise condition was more than twice the subtitle fixation duration in the noise condition. However, the effect was in the expected direction.

Finally, hypothesis 5b stated when there are subtitles in a video, attention would be higher in a video containing a person in an area without noise than when the video is played in area with noise. The average amount of subtitle fixation duration for participants in the no noise

condition was 0.3 seconds (*SD* = 0.5); that was shorter than the average time in the noise condition (*M* = 0.5, *SD* = 0.5). The independent sample *t* test indicated that this was not a statistically significant difference among the groups, $t(21) = -0.61$, $p = .55$. This signified that there was no statistical evidence in support of hypothesis 5b.

These two mean comparisons did not control for the other factors—person and or noise. An ANCOVA including the two factors, as well as three control factors—environmental attitude scale, and environmental interest scale, and recycling—also indicated that there is no statistically significant role of person $F(1, 43) = 1.7$, p = .2 and noise F(1, 43) 0.97, p = .33 in predicting total subtitle fixation, with a small effect sizes for person ($\eta_p^2 = .05$) and noise ($\eta_p^2 = .03$). The other independent variables and covariates were not statistically significant. The ANCOVA result was similar to the tests mentioned above, however recycling had a moderate effect size ($\eta_p^2 = .06$). The interaction effect was not at a statistically significant level either, despite having a decent effect size ($\eta_p^2 = .05$).

Table 2

Total subtitles fixation duration

Source	*Type III Sum of Squares*	*df*	*Mean Square*	*F*	*Sig.*	*Partial Eta Squared*
Corrected Model	3.015a	6	.502	1.341	.264	.179
Intercept	.016	1	.016	0.42	.838	.001
Recycle	.827	1	.827	2.207	.146	.056
Environmental Attitude Scale	6.154E-8	1	6.154E-8	.000	1.000	.000
Environmental Interest Scale	.076	1	.076	.203	.655	.005
Person	.637	1	.637	1.699	.200	.044

Noise	.360	1	.360	.961	.333	.025
Person * Noise	.695	1	.695	1.854	.182	.048
Error	13.864	37	.375			
Total	27.200	44				
Corrected Total	16.879	43				

a. R Squared = .179 (Adjusted R^2 = .045)

Discussion and conclusion

Overall, the present study sought out to determine what viewers attend to on social media by providing a combination of a person, subtitles, and noise elements and how, if any, effect occurred. It also looked to understand if the attention comes from having some level of impression-relevant involvement paired with attitude, interest, and or recycling habits. Because little research has been conducted in this area, especially using eye-tracking technology, and the interaction between these items, this research serves as the beginning of future research. The study included a convenience sample with participants between the ages of 18–26 years from a midsize Midwest university.

This present study contributed to future research in the field of social media by investigating what kind of video content individuals give attention to. Being able to produce video content the consumers would prefer to see could increase efficiencies and profits for media companies, marketing information for organizations, which in turn could contribute to a more informed public. Furthermore, this research helps open the door for more research to focus on different social media platforms and topics, as well as trying to discern a more cognitive meaning as to why individuals attend to certain types of media.

The results within the study were mixed. However, there was adequate support for attention on subtitles, which showed further support for Ma et al. (2005) and Chen et al. (2003) when considering textual attention in a video. While other hypotheses were not statistically significant (H1a, H1c, H2a, H2b, H2d, H3, H4, H5a, and H5b); however, some were in the expected direction. The main findings of this study were able to provide organizations, companies, marketers, advertisers, and more with a start in information regarding what individuals attend to. The independent variables in this study—person, subtitles, and noise—are factors to take into considering when creating materials to be seen by consumers on a regular basis. It is important to note that this is simply just the beginning on what could be further research in terms of video attention on social media. Additionally, this research could be used to help determine what kind of elements are needed to garner the attention to social causes, such as being environmentally friendly and recycling. When organizations know what viewers are more likely to attend to, it could lead to more social change, and a more informed public, which is needed, as suggested by Kumar and Gupta (2016).

As noted in the results section, attention was higher in a video when it contained subtitles compared to a video that does not contain subtitles, as stated by hypothesis 1b. This leads to future implications of research, which could discuss the cognitive effectiveness of subtitles on social media. For example, Perego et al. (2010) studied the cognitive effectiveness of subtitles in participants' after they were shown a video with no other distractions. While they found that there was recognition in subtitle and scene recognition (Perego et al., 2010), there could be a difference when social media is involved, because other posts by people can be considered distracting. It's important to note that the participants in the Perego et al. (2010) study were exposed only to the video content for 15 minutes and a survey following the video, and wore

headphones to minimize external noise and distraction. The cognitive effectiveness of videos with subtitles could prove to be beneficial information for companies and organizations trying to get their message across, whether it is about a specific product or a social issue.

An additional result supported hypothesis 2c, which stated, attention would be higher when a video contains subtitles for an individual in an area without noise than when there are no subtitles. The *t* test revealed that attention was higher in an area without noise when there was no subtitles, which is opposite of the anticipated direction. This result is contradictory toward the result of hypothesis 1b, which stated that attention would be higher when there are subtitles. These opposing results warrant further research specifically focused on subtitles and how noise can affect an individual's attention.

Furthermore, the small sample size had an effect on the overall results. For example, in hypotheses 5a and 5b, if there were a higher sample size, there results likely could have been statistically significant, both for the main effects for person and noise and for the interaction for person and noise. Additionally, there could have been a main effect for the recycling covariate.

Other observations

After reviewing the recordings on Tobii Studio 3.4.8.1348, it was noted that participants were reading the text portion of the post paired with the video. Interestingly enough, some of those that read the post appeared decided to scroll past the video, while a few decided to view the video (see Figure 2, Appendix E). These implications could lead to future research regarding the tone and contents message with a video, as well as the topic, which will be further discussed later in this paper.

An additional observation was the amount of participants that recycle. Previously mentioned, of all the 95 participants, 85% indicated that they recycle (*SD* = 0.36). Of those that

indicated they do not recycle, 14 participants, half of those responded that they had not been introduced to recycling in their academic career (elementary school, middle school, high school, or college). It was interesting to see that a majority of those who were introduced to recycling in their academic career, 78, still recycle at some level currently, of the 81 that indicated they recycle. Which means that only three participants that were not introduced to recycling earlier in their academic career recycle. This shows how the education system and its practices had an affect on its students and their futures, as well as for the future of society.

Limitations

One limitation could be using I-VT because of problems noted by (Salvucci & Goldberg, 2000). Such as when the point velocities are near the threshold—midway between saccades and fixation—it can result in "blips" in the identification (Salvucci & Goldberg, 2000). The problem can be prevalent when analyzing time-averaged data (Salvucci & Goldberg, 2000)

A limitation within the experiment was that the location is focused on the Midwest for a convenience sample. The survey used before the experiment is a limitation because of the self-reported nature of a survey and respondents could provide inaccurate data. Because the video content used will be determined by an earlier survey, this experiment had limitations of personal preferences by the participant. To better understand why users continued to watch a piece of video content, it is suggested that this current research piece could be used with qualitative research practices. Moreover, in this study, as Petty and Cacioppo (1986) discussed, there was "a natural confounding" variable, considering an individual's personal relevance with an issue, as well as the amount of thinking a person had previously done before participation in the experiment.

An additional limitation within this study was the video stimulus that was used. Rather than the researcher creating video stimulus to be used in the duration of the study, the researcher sought out video content that was already made and that could be edited it. This was a limitation because the researcher did not have total control over the video stimulus being used by seeking out video that met requirements rather than creating stimulus specifically for the study. Moreover, the video was conveyed in the positive tone, rather than having both a negative and positive tone video for participants.

Additional limitations include exclusivity regarding the devices and platform used in the study. The present research was performed on a desktop computer only and designed by focusing on Facebook. Furthermore, the profile on the Facebook design of was that of a woman, rather than a gender-neutral name. The profile image that was also used was not a gender-neutral appearing image. Nearly 50 out of 95 participants appeared to have looked at the name on the profile at the top of the page at one point to create a saccade. While there was no question directly asking if participants remembered or noticed the name, it is important to note that it was a woman and not a gender-neutral name and photo.

Future research

Because of the aforementioned limitations, this would allow for multiple areas of future research. Some of which would include looking specifically at other platforms, for examples Twitter and Instagram. This could also allow for additional research using a Facebook platform. Furthermore, there was the ability perform a study on different devices, rather than a desktop, like a phone, tablet, or laptop. One consideration for performing it on a different device is how the platforms are designed differently and have a different level of interactivity.

Additionally areas of future research could include a wide range of topics that do not focus primarily on environmentally friendly issues. Future topics could include politics, gender equality, climate change, sports, and the list could go on. If there is an issue that has been prevalent on social media, it could also be studied in the context of attention to what could be shown to participants. Future research could also look at the tone of the video, positive compared to negative. As Baumeister et al. (2001) wrote, the negative can be stronger than the good in some topics and aspects, such as health, social support, and learning topics. Alternatively, as discussed by Baumeister et al. (2001), and in the literature review, positive can be beneficial as the positive may predominate over the negative, especially in some psychological cases. Previously discussed, Baumeister et al. (2001) said if an individual wants to avoid a negative outcome in the future, positive messaging could be more effective. Because of the adverse effects, it is important to consider both negative and positive messaging in future research. In addition, this could lead to more cognitive research to have a greater understanding of how individuals perceive both messaging types, as well as the topic. Moreover, the text portion of the post with these topics should also be studied. The current research presented does not focus primarily on the text portion of the post, and doing so in future research would provide valuable insight. This insight could discuss positive and negative sentiment within a post and post length, as well as other possibilities.

Supplementary qualitative research could also be added to future research to better understand why a user may be interested in a video. There could be questions that ask why participants decided to continue watching a video they saw on social media, or why they scrolled past it. Information such as this would prove valuable to understanding different reasons why people might attend to videos, as well as opening the door for more future research. Overall, this

study could be considered part of the beginning to a better understanding of how consumers view videos on social media.

References

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323–370.

Bohlen, G., Schlegelmilch, B. B., & Diamantopoulos, A. (1993). Measuring ecological concern: a multi-construct perspective. *Journal of Marketing Management, 9*(4), 415–430.

Bojko, A. A. (2009). *Informative or misleading? Heatmaps deconstructed.* Paper presented at the International Conference on Human-Computer Interaction, Springer, Berlin, Heidelberg.

Briggs, S. R., Cheek, J. M., & Buss, A. H. (1980). An analysis of the self-monitoring scale. *Journal of Personality and Social Psychology, 38*(4), 679–686.

Broadbent, D. E. (1971). *Decision and stress*. Oxford, England: Academic Press.

Buttel, F. H. (1979). Age and environmental concern: A multivariate analysis. *Youth & Society, 10*(3), 237–256.

Center, P. R. (2018). *Social Media Fact Sheet*. Retrieved from http://www.pewinternet.org/fact-sheet/social-media/.

Chen, L.-Q., Xie, X., Fan, X., Ma, W.-Y., Zhang, H.-J., & Zhou, H.-Q. (2003). A visual attention model for adapting images on small displays. *Multimedia systems, 9*(4), 353–364.

Cho, H., & Boster, F. J. (2005). Development and validation of value-, outcome-, and impression-relevant involvement scales. *Communication Research, 32*(2), 235–264.

Cooper, B. B. (2013). 7 Powerful Facebook Statistics You Should Know About. Retrieved from https://www.fastcompany.com/3022301/7-powerful-facebook-statistics-you-should-know-about.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193–222.

Fishman, E. (2016). How Long Should Your Next Video Be? Retrieved from https://wistia.com/learn/marketing/optimal-video-length.

Hughes, D. J., Rowe, M., Batey, M., & Lee, A. (2012). A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage. *Computers in Human Behavior, 28*(2), 561–569.

Izard, C. E., & Ackerman, B. P. (2000). Motiviational, organizational, and regulatory functions of discrete emotions. In M. L. J. M. Haviland-Jones (Ed.), *Handbook of emotions* (2nd ed., pp. 253–264). New York: Guilford.

Jellinek, R. (2011). What is B-Roll and Why is it Valuable? Retrieved from https://contentmarketinginstitute.com/2011/10/what-is-b-roll-and-why-is-it-so-valuable/.

Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology, 37*(1), 43–75.

Johnston, W. A., & Heinz, S. P. (1978). Flexibility and capacity demands of attention. *Journal of Experimental Psychology: General, 107*(4), 420–435.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons, 53*(1), 59–68.

Kumar, V., & Gupta, S. (2016). Conceptualizing the evolution and future of advertising. *Journal of Advertising, 45*(3), 302–317.

Lai, K.-H., Cheng, T., & Tang, A. K. (2010). Green retailing: factors for success. *California Management Review, 52*(2), 6–31.

Leippe, M. R., & Elkin, R. A. (1987). When motives clash: Issue involvement and response involvement as determinants of persuasion. *Journal of Personality and Social Psychology, 52*(2), 269–278.

Ma, Y.-F., Hua, X.-S., Lu, L., & Zhang, H.-J. (2005). A generic framework of user attention model and its application in video summarization. *IEEE transactions on multimedia, 7*(5), 907–919.

O'Donnell, F. (2013). *Marketing to the Green Consumer*. Retrieved from http://academic.mintel.com.proxy.bsu.edu/insight_zones/2/.

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology, 64*, 77–107.

Perego, E., Del Missier, F., Porta, M., & Mosconi, M. (2010). The cognitive effectiveness of subtitle processing. *Media Psychology, 13*(3), 243–272.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Communication and Persuasion* (pp. 1–24): Springer.

PollyPoly. (2017, August 22). Times Square on rainy night. 3D audio. NYC Ambient scenes. G8512-60mm f3.5. Retrieved from https://www.youtube.com/watch?reload=9&v=xG3VRtk8xNw.

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological review, 78*(5), 391–408.

Salvucci, D. D., & Goldberg, J. H. (2000). *Identifying fixations and saccades in eye-tracking protocols.* Paper presented at the Proceedings of the 2000 Symposium on Eye Tracking Research & Applications, Palm Beach Gardens, Florida, USA.

Schlegelmilch, B. B., Bohlen, G. M., & Diamantopoulos, A. (1996). The link between green purchasing decisions and measures of environmental consciousness. *European Journal of Marketing, 30*(5), 35–55.

Sherif, M., & Cantril, H. (1947). *The psychology of ego-involvements: Social attitudes and identifications*. Hoboken, NJ, US: John Wiley & Sons Inc.

Silvia, P. J. (2006). *Exploring the Psychology of Interest*: Oxford University Press.

Smith, A., & Anderson, M. (2018). *Social Media Use in 2018*. Retrieved from http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/.

Smith, K. T. (2010). An examination of marketing techniques that influence Millennials' perceptions of whether a product is environmentally friendly. *Journal of Strategic Marketing, 18*(6), 437–450.

Stroud, N. J. (2017). Attention as a Valuable Resource. *Political Communication, 34*(3), 479–489.

Tobii. (n.d.). Digging Into Areas of Interest. Retrieved from https://www.tobiipro.com/learn-and-support/learn/steps-in-an-eye-tracking-study/data/digging-into-areas-of-interest/.

Vachon, F., Hughes, R. W., & Jones, D. M. (2012). Broken expectations: violation of expectancies, not novelty, captures auditory attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(1), 164–177.

Vraga, E., Bode, L., & Troller-Renfree, S. (2016). Beyond self-reports: Using eye tracking to measure topic and style differences in attention to social media content. *Communication Methods and Measures, 10*(2-3), 149-164.

Warshaw, P. R. (1978). Application of selective attention theory to television advertising displays. *Journal of Applied Psychology, 63*(3), 366–372.

Webb, T. W., & Graziano, M. S. A. (2015). The attention schema theory: A mechanistic account of subjective awareness. *Frontiers in Psychology, 6*(1–11), 500.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*(2, Pt.2)), 1–27.

Zimbardo, P. G. (1960). Involvement and communication discrepancy as determinants of opinion conformity. *The Journal of Abnormal and Social Psychology, 60*(1), 86–94.

Appendix A

Video stimulus conditions as seen in the design

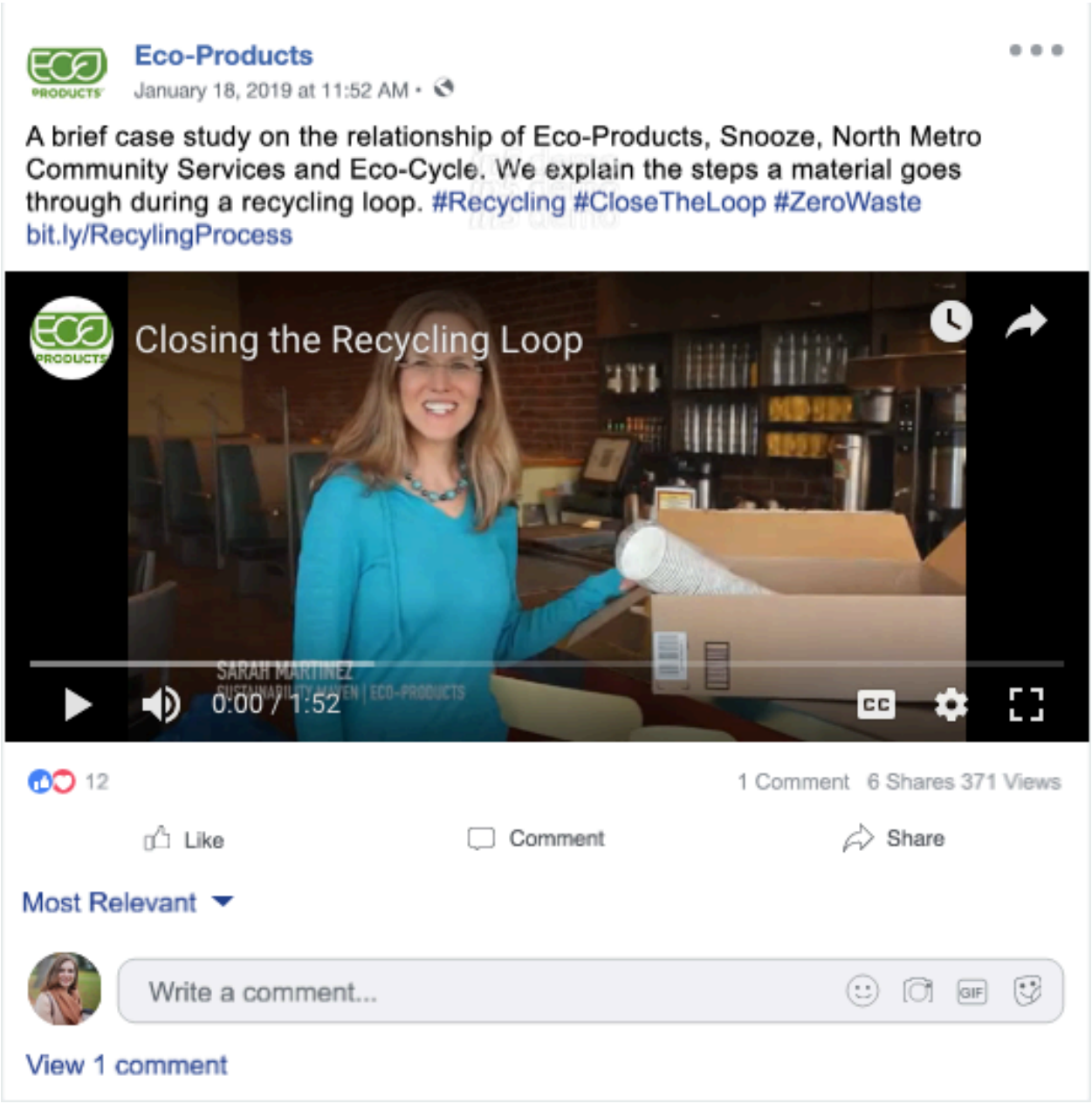
Condition one: No person, no subtitles

Note: This condition does not account for the noise, no noise variable.

Condition two: No person, subtitles

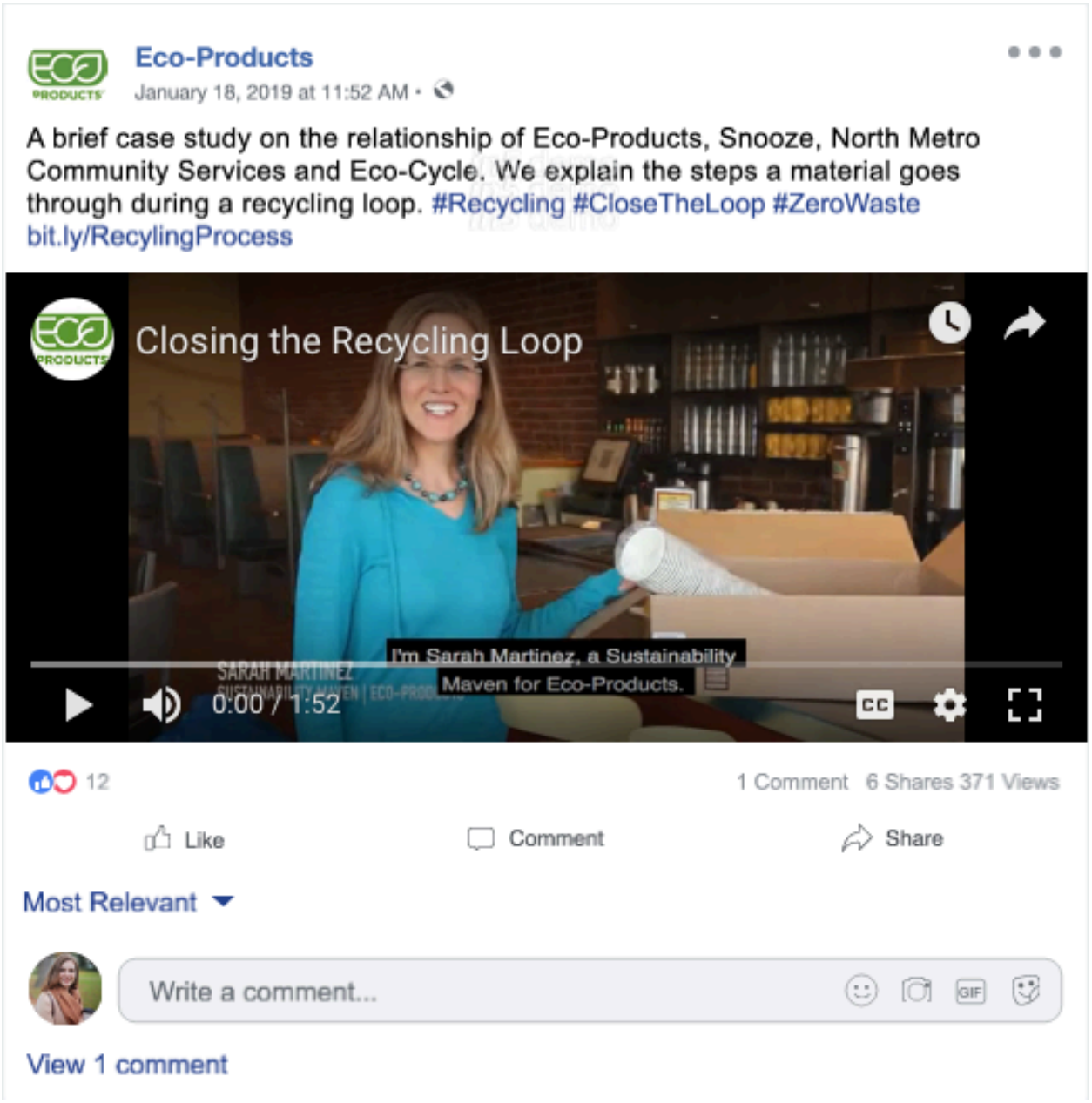
Note: This condition does not account for the noise, no noise variable.

Condition three: Person, no subtitles



Note: This condition does not account for the noise, no noise variable.

Condition four: Person, subtitles

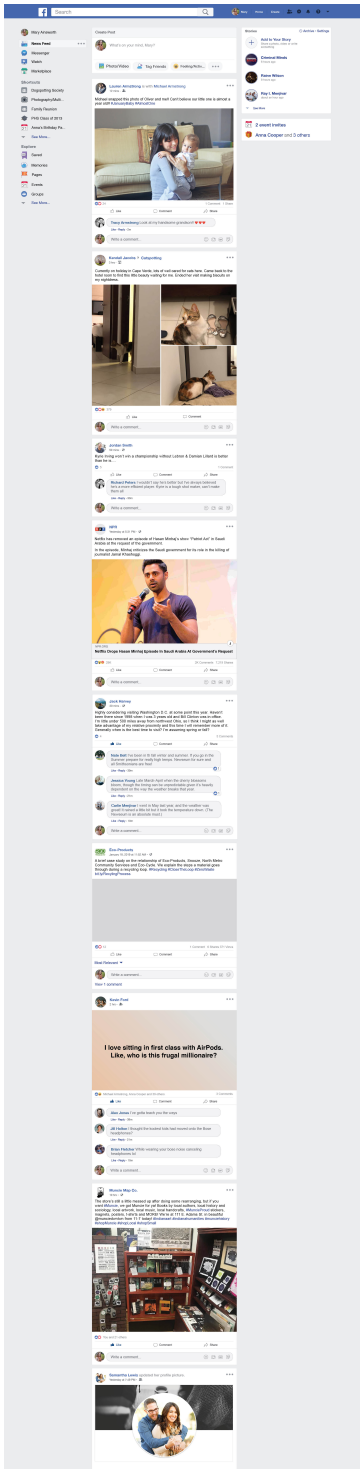


Note: This condition does not account for the noise, no noise variable.

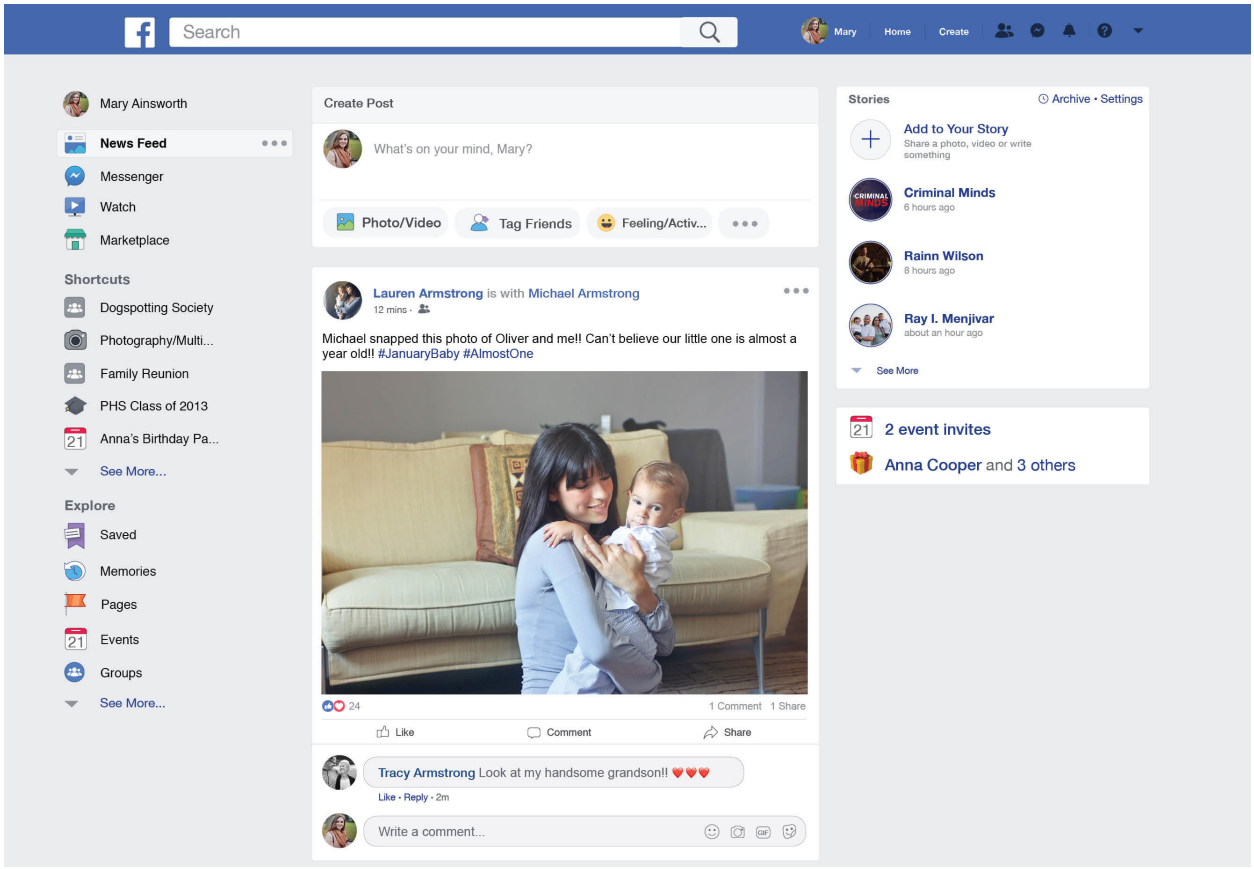
Appendix B

Facebook stimulus

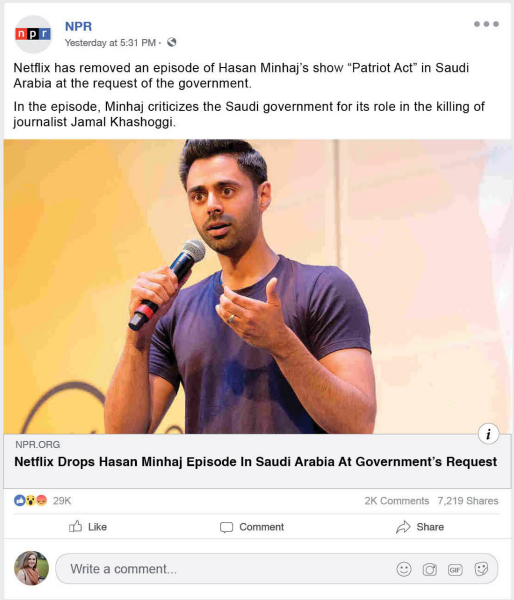
This is the overall design of the Facebook design created for this study.

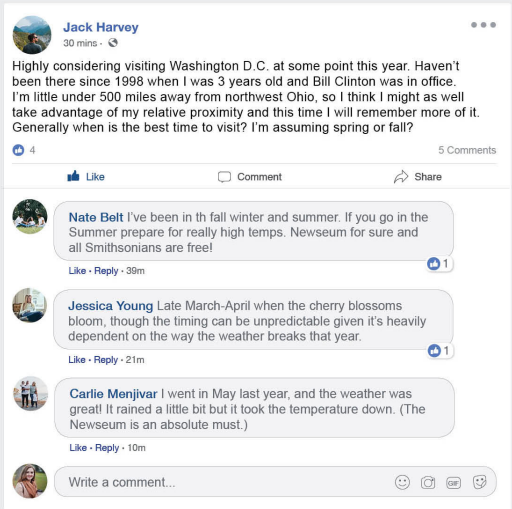


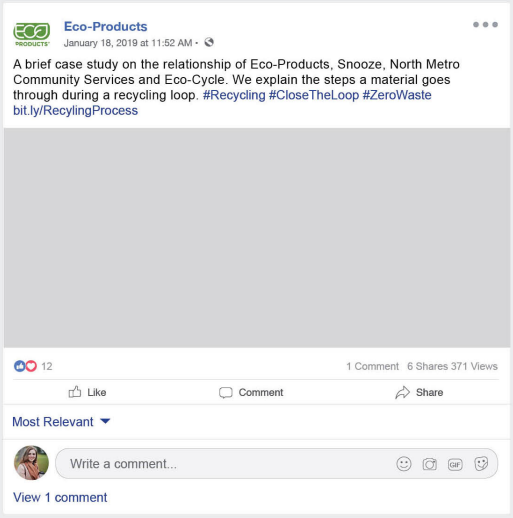
The following include more clear images of the posts.



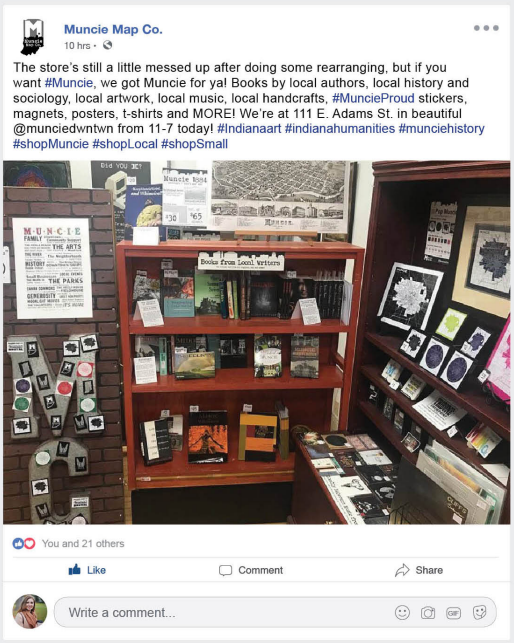

Jordan Smith
50 mins ·
Kyrie Irving won't win a championship without Lebron & Damian Lillard is better than he is....
5
1 Comment
Like
Comment
Share
Richard Peters I wouldn't say he's better but I've always believed he's a more efficient player. Kyrie is a tough shot maker, can't make them all
Like · Reply · 30m
Write a comment...

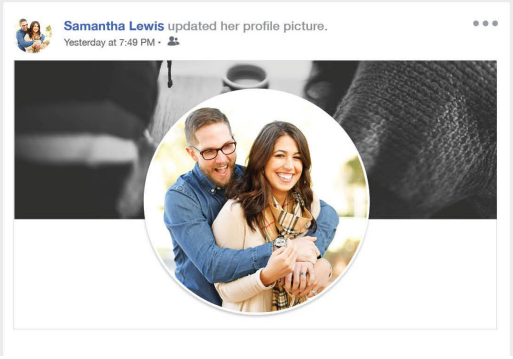

NPR
Yesterday at 5:31 PM ·
Netflix has removed an episode of Hasan Minhaj's show "Patriot Act" in Saudi Arabia at the request of the government.
In the episode, Minhaj criticizes the Saudi government for its role in the killing of journalist Jamal Khashoggi.
NPR.ORG
Netflix Drops Hasan Minhaj Episode In Saudi Arabia At Government's Request
29K
2K Comments 7,219 Shares
Like
Comment
Share
Write a comment...


Jack Harvey
30 mins ·
Highly considering visiting Washington D.C. at some point this year. Haven't been there since 1998 when I was 3 years old and Bill Clinton was in office. I'm little under 500 miles away from northwest Ohio, so I think I might as well take advantage of my relative proximity and this time I will remember more of it. Generally when is the best time to visit? I'm assuming spring or fall?
4
5 Comments
Like
Comment
Share
Nate Belt I've been in th fall winter and summer. If you go in the Summer prepare for really high temps. Newseum for sure and all Smithsonians are free!
Like · Reply · 39m
Jessica Young Late March-April when the cherry blossoms bloom, though the timing can be unpredictable given it's heavily dependent on the way the weather breaks that year.
Like · Reply · 21m
Carlie Menjivar I went in May last year, and the weather was great! It rained a little bit but it took the temperature down. (The Newseum is an absolute must.)
Like · Reply · 10m
Write a comment...



Note: The video was coded into the design using Brackets and was not originally designed into the layout because videos cannot be coded into designs on Adobe InDesign.


Muncie Map Co.
10 hrs
The store's still a little messed up after doing some rearranging, but if you want #Muncie, we got Muncie for ya! Books by local authors, local history and sociology, local artwork, local music, local handcrafts, #MunciePROUD stickers, magnets, posters, t-shirts and MORE! We're at 111 E. Adams St. in beautiful @munciedwntwn from 11-7 today! #Indianaart #indianahumanities #munciehistory #shopMuncie #shopLocal #shopSmall
Did You Be?
Books from Local Writers
You and 21 others
Like
Comment
Share
Write a comment...


Samantha Lewis updated her profile picture.
Yesterday at 7:49 PM

Appendix C

Survey for participants

Social media usability

Please respond to the following statements	Strongly Disagree	Disagree	Neither (Dis)Agree	Agree	Strongly Agree
I will implement recycling into my lifestyle.	O	O	O	O	O
I believe recycling has a positive impact on the environment.	O	O	O	O	O
Family influences me about my recycling habits.	O	O	O	O	O
Friends influence me about my recycling habits.	O	O	O	O	O
Celebrities and or social media influencers affect my recycling habits.	O	O	O	O	O
Incorporating recycling into my lifestyle is difficult.	O	O	O	O	O
Recycling services, for the home, are expensive.	O	O	O	O	O
Recycling services, for the campus, are expensive.	O	O	O	O	O
I know how to properly recycle.	O	O	O	O	O
Recycling bins are easily accessible.	O	O	O	O	O
The environment is one of the most important issues facing society today.	O	O	O	O	O
We should pay a considerable amount of money to preserve our environment.	O	O	O	O	O
Strict global measures must be taken immediately to halt environmental decline.	O	O	O	O	O
A substantial amount of money should be devoted to environmental protection.	O	O	O	O	O

Unless each of us recognizes the need to protect the environment, future generations will suffer the consequences.	O	O	O	O	O
The environmental policies of the main political parties are one issue 1 consider when deciding how to vote.	O	O	O	O	O
Personally, I cannot help to slow down environmental deterioration.	O	O	O	O	O
The importance of the environment is frequently exaggerated.	O	O	O	O	O
The benefits of overcoming environmental deterioration are not sufficient to warrant the expense involved.	O	O	O	O	O
Even if each of us contributed towards environmental protection, the combined effect would be negligible.	O	O	O	O	O

Please respond to the following statements	Strongly Disagree	Disagree	Neither (Dis)Agree	Agree	Strongly Agree
I actively recycle in my every day life.	O	O	O	O	O
I feel good when I recycle.	O	O	O	O	O
I feel motivated to recycle after viewing the Facebook newsfeed.	O	O	O	O	O
I feel motivated to recycle when I see trash in public.	O	O	O	O	O
I'm making the world a better place when I recycle.	O	O	O	O	O
I feel motivated to recycle when others around me recycle.	O	O	O	O	O

When did you become interested in recycling?
Elementary school
Middle school
High school
College
N/A

How old are you?
18
19
20
21
22
23
24
25
26+

What is your gender?
Female
Male
Prefer not to answer

Do you recycle?
Yes
No

Appendix D

Study procedure for the experiment

S T U D Y P R O C E D U R E

1. Recruitment

a. Students will be recruited from TCOM 100, TCOM 206, TCOM 345, JOUR 103, JOUR 241, JOUR 250, JOUR 354, and JOUR 456.

b. Students will be recruited in their classrooms at their meeting times (various times). The principal investigator (PI) will present a slideshow regarding information about the study, where it takes place, what benefits they will receive (2% extra credit), and how to schedule an appointment.

- **i.** Students will schedule themselves for an appointment via a Doodle Poll link provided during the class meeting time.

c. Consent forms will be signed when students arrive to their eye tracking appointment.

2. Study

a. Before the participant arrives, the PI will set up the appropriate stimuli and set the condition accordingly.

- **i.** The PI will pull the webpage for a condition—one, two, three, four—and answer a question before the calibration to signify which condition it is.
- **ii.** In the event of noise being played, the PI will answer the question to answer if noise will be played or not.
    1. On the PI's computer, the selected noise (Times Square, New York City, New York) will be ready to be played. It will begin playing when the participant leaves the instruction screen.
- **iii.** The PI will check off the box when it is completed.

b. The participants will arrive to AJ 386.

- **i.** The PI will invite them in and advise them to sit down in front of the computer.
    1. Then participants will receive the consent form to sign, which also ensures they receive extra credit for arriving to their appointment.
- **ii.** Depending on the participant, the PI will advise them to sit closer, adjust their chair, and to not look away from the screen after the calibration has started.

iii. The PI will explain the calibration process—there will be two lines on the screen and the participant should focus where they intersect, a small gray square.

1. If recalibration is necessary, the PI will notify the participant and it will be completed.

c. Participants will begin the eye-tracking portion of the study.

i. After the calibration is complete, participants will be shown the instructions. They will be then be notified to let the PI know when they are finished looking at the newsfeed, and to press spacebar to begin.

1. Participants will be asked to imagine they are in Times Square. If they will have noise, they will be notified that there will be sound to make it more realistic. The PI's MacBook Pro will be used with the volume on the computer up to 12 and the speakers connected to the computer are at the halfway point.

ii. Participants will scroll through Facebook feed at their leisure.

iii. Participants will notify the PI when they are done.

iv. Participants will be prompted with an instruction screen to answer statements using a Likert Scale and other questions.

v. Participants will answer the statements and questions that appear on the screen. At which point the PI will remove their self from the desk. This is to ensure the participants do not feel inclined to answer a certain way.

vii. When they have reached the end of the survey, they will be prompted with a debriefing slide.

viii. Participants will then be notified by the PI that they have completed the research, and thanked for their time.

3. Data collection

a. Data from the participants will automatically be collected using Tobii Studio.

b. Data will be stored on a password-protected computer in a locked room, indefinitely.

c. Data will be downloaded from Tobii periodically to ensure it will be not be lost.

d. In the event of a screen recording issue during data collection, the data will still be collected. The PI will watch the recordings afterward and determine the AOIs based off the recordings.

i. AOIs must be made after the recording is complete.

ii. The PI will draw the AOIs in Tobii, which will then be used to determine how long individuals looked at the content.

Appendix E

Heat maps

Figure 1.

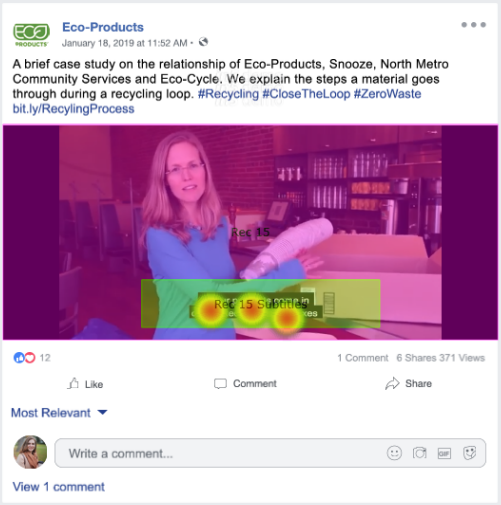


Figure 1 represents the heat map recording from the person, subtitles, and no noise condition by a participant.

Figure 2.

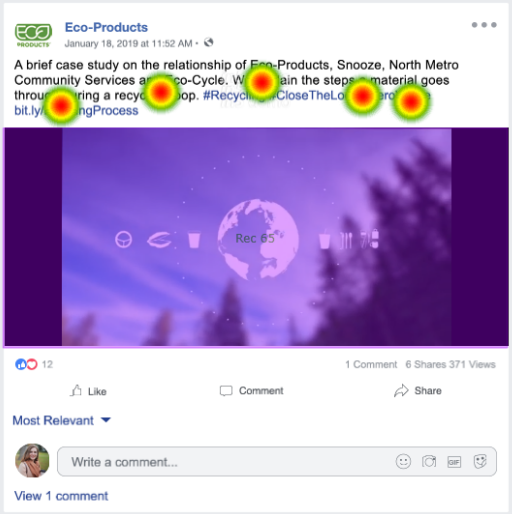


Figure 2 represents the heat map from a participant in the no person, no subtitles, and no noise condition.